Stochastic Hyperbolic Discounting: An Application to an Environmental Dynamic Game

Luis-Javier Capsi Morales

MSc in Economics, Universitat de Barcelona

Advisors: Jesús Marín-Solano and Jorge Navas

June 2018

Abstract

We study the infinite horizon emissions and stock of pollution choices of time-inconsistent individuals by incorporating the stochastic hyperbolic preferences of Harris and Laibson (2013), later extended by Zou et al. (2014), into the environmental dynamic game proposed by Jørgensen et al. (2003) with linear-state structure. We derive analytic solutions for optimal emissions and stock of pollution selections for sophisticated individuals and extend the results with a sensitivity analysis of the stochastic hyperbolic parameters and their impact on the economy. Compared to the results of Jørgensen et al., we find that the stochastic hyperbolic discounting model increases the emissions rates and the stock of pollution, concretely, in the case in which agents are highly impatient.

Keywords: Stochastic hyperbolic discounting, time-inconsistency, infinite lifetime, dynamic games, environmental economies, cooperation, sophisticated agent.

1 Introduction

Economic decisions are, in most cases, of the intertemporal kind and imply trade-offs between present and future costs and benefits. Since Samuelson (1937), the Discounted Utility Model, also known as the Exponential Discounted Utility Model with a constant discount rate, became the mainstream in this framework. However, Strotz (1956) found that this was the only discount function that lead to time-consistent preferences, i.e., preferences where individuals do not have incentives to change their ex-ante optimal plan when it is recomputed in the future.

Evidence from psychology and behavioral economics has shown that the inconsistent behavior is common in human time preferences (Thaler, 1981; Thaler and Shefrin, 1981; Ainslie, 1992; Loewenstein and Prelec, 1992; Laibson, 1997; McClure et al., 2004). Those studies have been evolving rapidly, from Thaler (1981), where the author notices that individuals are discounting time differently depending, on the length of time, size of the reward and the sign, i.e., different discounting rates for gains and losses; to Laibson (1997), who explains how people fight this time-inconsistency by obtaining less liquid assets. More recent studies have shown that some individuals changed their ex-ante choices when they were asked again in the future and that people are willing to incorporate reminders on their portfolio activities to not only do not withdraw their accounts, but also keep depositing more savings in the future (Havely, 2015; Karlan et al., 2016).

The non-constant discounting models in general and, in particular, the Hyperbolic Discounting, which addresses the evidence by giving different weights to present and future ($\beta - \delta$ preferences), have become a major topic in intertemporal choice models. Works, such as Barro (1999), who adapts the classic Ramsey model to the non-constant discounting, or O'Donoghue and Rabin (1999), in which the authors show how agents with hyperbolic preferences might face self-control issues, are examples of the vast literature that has been developed. In fact, hyperbolic discounting assumes that the capability to compute problems, and hence, make distinctions, diminishes for more distant events, as Thaler (1981) illustrates on his dynamic inconsistency example, which stresses that the elapse of small periods of time in the long future are nearly irrelevant. This concept is key when considering long-lived problems of the environment, where the present agents might feel substantially closer to their children, than to their grandchildren or even further generations, so that, making current time agents be willing to discount the welfare of next generations (Karp, 2005).

Three behavior profiles for agents are found in the hyperbolic discounting framework, which are separated according to their ”consciousness of their inconsistency”. That is, whether they are aware of their time-inconsistency or not. *Naive-agents* do not know about their inconsistency; in which case, they will change their strategies in the future repeatedly. The case of *precommited-agents* considers a scenario in which agents are aware of their inconsistent preferences and they can fix a binding contract, prohibiting them from deviating from their ex-ante choices. Otherwise, if they are aware of this but they cannot enforce this contract, *sophisticated-agents* will have completely rational strategies, since they will advance their future deviations and incorporate this information in the choice process.

With regards to modelling pollution problems, three characteristics must be considered: *interdependence*, when the actions of individuals do not only affect their own utility, but also the utility of other agents; *time*, since environmental problems are intrinsically dynamic; *strategic behavior*, strategies of other agents affect the final output. Jørgensen et al. (2003) identified agreeability conditions for linear-state differential games in which time consistency was found, which is necessary to achieve dynamic (intertemporal) individual rationality. In the case in which each agent's cooperative payoff dominates her non-cooperative payoff at any instant of time, the dynamic individual rationality is ensured. The authors consider the association between several agents – two countries – which have to ratify a cooperation agreement to control of the transboundary pollution for infinite periods ($T \to \infty$). This is extended from Petrosjan and Zaccour (2003), in which they obtain a time-consistent Shapley value allocation for pollution costs in the control, and earlier transboundary pollution models, such as van der Ploeg and de Zeeuw (1992).

In this paper, we study the intertemporal emissions rules and stock of pollution for economies, which pollute as a consequence of the production of consumption goods, focusing on sophisticated individuals with time-inconsistent preferences and an infinite lifetime. Almost all existing studies on intertemporal pollution problems assume the scenario of the classical exponential discounted utility model. By considering hyperbolic discounting, one country can be seen as the sum of infinite future different countries, in which case, we address also the problem of "internal" time-inconsistency. We contribute to the existing literature by incorporating the model of stochastic hyperbolic discounting, developed by Harris and Laibson (2013) and extended by Zou et al. (2014), into the differential game for pollutant economies developed by Jørgensen et al. (2003). We find time-consistent policies for three different cases: *single-agent* case and, for two countries, the *non-cooperative (Nash)* and *cooperative* cases. By comparing with the results of Jørgensen et al. (2003), we find that emissions rates and stocks of pollution increase when the discount function is described by stochastic hyperbolic discounting.

Our paper complements the study by showing several approximations for the two stochastic hyperbolic discounting preferences' main parameters. We compare how emissions rates and pollution stocks evolve for the three different agent's cases, considering different time preferences: the standard case (exponential discounting utility model), the Instantaneous Gratification (IG) case (Harris and Laibson, 2013) and intermediate cases.

The rest of the paper is organized as follows. Section 2 reviews stochastic hyperbolic discounting and the pollution differential game. Section 3 introduces the basic model, which incorporates stochastic hyperbolic discounting into the pollution differential game, setting a general model for n-agents. Section 4 describes and solves the dynamic programming equations for the three particular cases. Section 5 presents and discusses the sensitivity analysis for the stochastic hyperbolic discounting parameters. Section 6 concludes and the appendices provide all the technical computations.

2 Preliminaries

In this section, we first introduce the modeling framework of Jørgesen et al. (2003) for pollutant economies. After this, we discuss the literature regarding the stochastic hyperbolic discounting preferences of Harris and Laibson (2013) and how Zou et al. (2014) incorporate this in their model.

2.1 Emissions and stock of pollution

We consider the scenario of two countries, which are subjected to pollution if they want to produce consumption goods, and their polluting emissions are accumulating in a common transboundary stock of pollution, e.g., (van der Ploeg and de Zeeuw, 1992). In addition to this, due to the nature of the stochastic hyperbolic discounting, we also consider the simplest case for one country, which sees the individual decision making as a dynamic game between the same agent in different points of time. Each country suffers welfare damage from the transboundary stock of pollution and produces a single consumption good Q_i. The production of such consumption good implies the emission of X_i pollution units, where $X_i = \varepsilon Q_i$. The parameter ε gives the ratio of emissions per output unit. For simplicity and without loss of generality we suppose $\varepsilon = 1$.

The evolution of the stock of pollution $\dot{P}(s)$ is represented by the following differential equation

$$\dot{P}(s) = \sum_{i=1}^{n} \omega_i X_i(s) - \delta P(s), \quad P(t) = P_t \geq 0, \tag{1}$$

where emissions X_i are transformed by a positive parameter ω_i into stock of pollution, which can be seen as environmental-oriented technology of country i, δ is the natural absorption rate and P_t is the value of the stock at the moment of the computation.

The welfare function is divided into two parts. On the one hand, instantaneous benefits from consumption for country i come from an increasing and concave function $U_i(Q_i)$. On the other hand, an increasing and convex function $D_i(P)$ incorporates the damage dealt by pollution. Hence, the welfare function for country i is described by

$$W_i(Q_i, P) = U_i(Q_i) - D_i(P)\,.$$

Given that $\varepsilon = 1$ and the relation between output and emissions, country's i aim is to find the emissions control strategy $X_i(s)$ which maximizes the welfare function subject to the stock of pollution motion, equation (1). Therefore, the dynamic problem for each country at time t can be expressed as

$$\max_{X_i} \int_t^{\infty} e^{-\rho(s-t)} W_i(X_i, P)\, ds,$$

$$s.t. \quad \dot{P} = \sum_{i=1}^{n} \omega_i X_i - \delta P, \quad P(t) = P_t, \quad i = 1, 2, ..., n\,;$$

with a state variable $P(s)$ - the stock of pollution - and each player's control variable X_i, i.e., pollutant emissions. In addition, ρ is a constant time discount rate. It might be interpreted to be equal to the risk-free interest rate.

Owing to the specifications from Jørgensen et al. (2003) for tractable differential games where, due to stationarity, the equilibrium strategies and the value functions do not depend explicitly on time, the particular functional forms for $U_i(X_i)$ and $D_i(P)$ for the linear-state differential game are

$$U_i(X_i) = \gamma_i \ln(\alpha_i X_i), \quad D_i(P) = \varphi_i P, \quad \gamma_i, \alpha_i, \varphi_i > 0, \tag{2}$$

where γ_i and α_i are parameters which represent country's i overall technology, and φ_i is a parameter that reflects the socio-economic sensibility and/or a direct damage to humans' health.

2.2 Stochastic hyperbolic discounting

Zou et al. (2014) aimed to study the consumption and portfolio problem of an individual with time-inconsistent preferences. They incorporated the stochastic hyperbolic discounting function of Harris and Laibson (2013) into Merton's classical framework (1969, 1971). Moreover, to obtain a time-consistent policy, the authors used the procedure followed by Marín-Solano and Navas (2009, 2010), converting the continuous time problem into a discrete time setting and solving it by backward induction. To do so, they obtained the agent's last future self optimal actions and then, by iteration and passing to the continuous time limit, they were able to find the dynamic programming equation for the sophisticated individual.

Based on Harris and Laibson (2013), the discount interval was divided into two subintervals: present and future. The present benefits were discounted exponentially with a constant discount rate ρ, whereas future benefits are further discounted by a factor β, where $0 < \beta \leq 1$, which epitomizes the agent's bias to the present or impatience. According to that, the discount function $\theta(t, s)$ is expressed as

$$\theta(t,s) = \begin{cases} e^{-\rho(s-t)}, & s \in [t, t+\tau), \\ \beta e^{-\rho(s-t)}, & s \in [t+\tau, \infty), \end{cases} \tag{3}$$

where $[t, t+\tau)$ is the present interval and $[t+\tau, \infty)$ is the future interval. The stochastic hyperbolic discounting function $\theta(t, s)$ satisfies stationarity by definition, which means that $\theta(t, t+s) = \theta(0, s)$. The present interval duration τ will be stochastic and exponentially distributed by a hazard rate parameter λ, with the expected duration $E[\tau] = 1/\lambda$. It implies that, the smaller λ is, the larger is the expected duration of the present interval τ. Interestingly, when $\lambda = 0$, the present interval's duration τ is equal to ∞, so that the problem becomes the general case of an exponential with a constant discount rate ρ. A similar result is provided by the case where $\beta = 1$, therefore, the value of λ is irrelevant. In the case in which $\lambda \to \infty$, τ will be equal to 0, accordingly, the discount function becomes an Instantaneous Gratification (IG) discounting function

$$\theta(t,s) = \begin{cases} 1, & s = t, \\ \beta e^{-\rho(s-t)}, & s \in (t, \infty), \end{cases}$$

in which the stochastic discount function is a deterministic jump function with a jump at t. Otherwise, when $0 < \lambda < \infty$, the problem becomes an interesting intermediate and more realistic case which allows us to make a sensitivity analysis.

3 Basic Model

In this section, we merge both settings previously discussed. We first set the stochastic hyperbolic discounting into the pollution linear-state differential game for n-agents. After that, we write the dynamic programming equation derived by Zou et al. (2014), for our setting. Finally, we discuss the proposed value functions for such differential game.

Considering now the nature of the stochastic hyperbolic discounting function $\theta(t, s)$, the dynamic problem for each country at time t have to be separated in two subintervals. Taking into account that each country i can be experiencing a different value for λ, we must pay heed to $E[\tau_i] = 1/\lambda_i$. Given these circumstances, we adapt the dynamic problem to equation (3), which leaves the general problem

$$\max_{X_i} E\left[\int_t^{t+\tau_i} e^{-\rho(s-t)} W_i(X_i, P)\, ds + \beta \int_{t+\tau_i}^{\infty} e^{-\rho(s-t)} W_i(X_i, P)\, ds\right], \tag{4}$$

$$s.t. \quad \dot{P} = \sum_{i=1}^{n} \omega_i X_i - \delta P, \quad P(t) = P_t\,, \quad i = 1, 2, ..., n\,. \tag{5}$$

Now, to obtain time-consistent policies, sophisticated agents with inconsistent preferences consider their future selves' preferences at their current decision making problem. By building the dynamic programming equation for sophisticated individuals, Zou et al. (2014) were allowed to find such policies and delete the uncertainty, giving an exact value for each country's τ_i. They computed the problem for value functions V_i, independent of time s, only depending on the current conditions of the state variable P. Therefore, our dynamic programming equation is articulated by

$$\rho V_i(P) + K_i(P) = \max_{X_i \geq 0} \left[W_i(X_i, P) + V_i'(P) \cdot g\Big(X_i, P\Big)\right], \tag{6}$$

where

$$K_i(P) = \lambda_i(1 - \beta_i) \int_t^{\infty} e^{-(\lambda_i+\rho)(s-t)} W_i(X_i^*, P)\, ds, \quad i = 1, 2, ..., n\,. \tag{7}$$

Note that equation (6) must be computed for every agent, in which case it becomes a system of dynamic programming equations. Denoted by $V_i(P)$ we have the value function for player i, with its derivative with respect to P, $V_i'(P)$. Such derived value function multiplies the function $g(X_i, P)$, which describes the equation of motion $\dot{P}$, $dP/dt = \dot{P} = g(X_i, P)$.

After maximizing the right hand side of equation (6) and substitute it inside equation (7), we will be able to split equation (6) into terms which multiply P(t) and otherwise, allowing us to find the consistent results. Notably, when $\lambda_i = 0$ or $\beta_i = 1$, i.e., the general case, the right hand side of equation (7) also becomes equal to 0.

Our guessing for the value functions $V_i(P)$ are the same as Jørgensen et al. (2003) proposed, since the form of the welfare functions $W_i(X_i, P)$ is also the same. According to this, the following expression denotes our proposed value functions for n countries

$$V_i\left(P\right) = A_i P + B_i\,, \quad i = 1, 2, \dots, n; \tag{8}$$

deriving with respect to P,

$$V_i^{'}\left(P\right) = A_i\,, \quad i = 1, 2, \dots, n; \tag{9}$$

where A_i and B_i are functions which are assumed to be independent on time s. Notably, due to our studied functional forms, neither the efficient emissions X_i, nor the stock of pollution trajectory $P(s)$, depend on B_i.

4 Time-consistent policies

Our analysis is focused on three different cases. We first discuss and solve the single-agent problem, which allows us to generate general conclusions for problems of this kind, and also to generate graphs which are highly useful for Section 5. After that, we discuss and solve the non-cooperative and cooperative cases for two players, enabling us to analyze and make comparisons between them.

4.1 Single-agent case

In the single-agent case, the economy is described by a unique country, which does not need to take into account other countries' emissions on the stock of pollution. Under the assumption of standard discounting, considering the single-agent case would of be poor relevance. However, such country must take into account its future country selves' preferences in its current decision-making. The more such country pollutes today, the less it might be able to pollute in the future to remain in the same stock of pollution. Now, adjusting equations (4-5) to the single-agent case and our particular functional form from equation (2), $\gamma \ln(\alpha X) - \varphi P$, the dynamic net payoff of emissions problem to be maximized is described by

$$J = E\Bigg[\int_t^{t+\tau} e^{-\rho(s-t)}\Big(\gamma \ln(\alpha X) - \varphi P\Big)\, ds + \beta \int_{t+\tau}^{\infty} e^{-\rho(s-t)}\Big(\gamma \ln(\alpha X) - \varphi P\Big)\, ds\Bigg], \quad (10)$$

$$s.t. \quad \dot{P} = \omega X - \delta P, \quad P(t) = P_t. \quad (11)$$

Thus, following the procedure for the sophisticated agent with stochastic hyperbolic discounting, the dynamic programming equation to be solved is

$$\rho V(P) + K(P) = \max_{X \geq 0}\Bigg[\Big(\gamma \ln(\alpha X) - \varphi P\Big) + V'(\omega X - \delta P)\Bigg], \quad (12)$$

where

$$K(P) = \lambda(1-\beta)\int_t^{\infty} e^{-(\lambda+\rho)(s-t)}\Big(\gamma \ln(\alpha X^*) - \varphi P\Big)\, ds. \quad (13)$$

We enclose all the calculations for the single agent case in the Appendix A with the aim of making the reading easier and comfortable.

By maximizing the right hand side of the equation (12) we find that the optimal emissions rule, equation (14), does not depend directly on time, but depends on A, i.e., the derivative of our proposed value function, equation (9). Then,

$$X^* = -\frac{\gamma}{\omega V'} = -\frac{\gamma}{\omega A}, \quad (14)$$

in which

$$A = \frac{\varphi}{\rho+\delta} \times \Big[\Omega - 1\Big], \quad \Omega(\lambda, \beta),$$

with the following term

$$\Omega = \frac{\lambda(1-\beta)}{\lambda+\rho+\delta}. \quad (15)$$

Note that Ω does not depend on time, which coincides with our previous assumption, making A constant in terms of time s. Such Ω has to be seen as the connector between λ cases. In the case in which $\lambda = 0$ or $\beta = 1$, the standard case, the whole term Ω becomes equal to 0, leaving the same result as in the standard case. Then, if $\lambda \to \infty$, known as the Instantaneous Gratification (IG) case, Ω will be reduced to $1 - \beta$. Otherwise, it allows the parameters ρ and δ to have an impact through Ω.

Proposition 1. Using the sophisticated agent procedure with time-inconsistent preferences, we find the following time-consistent and constant emissions rate and stock of pollution trajectory.

The time-consistent emissions rule follows as

$$X^* = -\frac{\gamma(\rho+\delta)}{\omega\varphi[\Omega - 1]}. \tag{16}$$

Note that in the standard case $\lambda = 0$ so that, Ω equal to 0, the optimal emissions rule X^* is positive. Also for the Instantaneous Gratification (IG) case $\lambda \to \infty$, thus, Ω equal to $1 - \beta$, the optimal emissions rule is positive and larger due to $0 < \beta \leq 1$.

Furthermore, the time-consistent stock of pollution trajectory,

$$P(s) = \left[P_t - \frac{(\rho+\delta)\gamma}{\delta\varphi} \times \frac{-1}{\Omega - 1}\right] e^{-\delta(s-t)} + \frac{(\rho+\delta)\gamma}{\delta\varphi} \times \frac{-1}{\Omega - 1}\,, \tag{17}$$

which in case $\lambda = 0$, the multipliers cancel and we obtain the standard case, i.e., exponential discounting with a constant discount rate, such as

$$P(s) = \left[P_t - \frac{(\rho+\delta)\gamma}{\delta\varphi}\right] e^{-\delta(s-t)} + \frac{(\rho+\delta)\gamma}{\delta\varphi}.$$

Interestingly, for the instantaneous gratification case, instead of cancelling the multipliers from equation (17), they become $1/\beta$, increasing the term $\frac{(\rho+\delta)\gamma}{\delta\varphi}$. Such behaviors are illustrated in figure (4) in Section 5.

4.2 Non-cooperative case (Nash equilibrium)

For the non-cooperative agents case, each country will maximize its own objective J_i^{nc} taking into account the emissions rules of the other player. Hence, the following dynamic net payoff functions to be maximized for non-cooperative agents with functional forms of the kind of equation (2) are given by

$$\begin{aligned} J_i^{nc} &= E\Bigg[\int_t^{t+\tau_i} e^{-\rho(s-t)}\Big(\gamma_i \ln(\alpha_i X_i) - \varphi_i P\Big)\, ds \\ &+ \beta \int_{t+\tau_i}^{\infty} e^{-\rho(s-t)}\Big(\gamma_i \ln(\alpha_i X_i) - \varphi_i P\Big)\, ds\Bigg], \end{aligned} \tag{18}$$

$$s.t. \quad \dot{P} = \omega_1 X_1 + \omega_2 X_2 - \delta P; \quad P(t) = P_t, \quad i = 1, 2. \tag{19}$$

Even though they only maximize their own function, i.e., inside J_i^{nc} only appears the i-agent's parameters, the stock of pollution shows the impact of having two polluting countries in the market. Note that we allow τ_i, the length of the current period, for the possibility to be different for each country, since they are computing their own J_i^{nc}, which implies that they might be experiencing a different λ_1, λ_2.

Again, under the standard procedure, we build the following dynamic programming equation for non-cooperative sophisticated individuals

$$\rho V_i^{nc}(P) + K_i(P) = \max_{X_i \geq 0}\Bigg[\Big(\gamma_i \ln(\alpha_i X_i) - \varphi_i P\Big) + V_i^{nc'}(\omega_1 X_1 + \omega_2 X_2 - \delta P)\Bigg], \tag{20}$$

where

$$K_i(P) = \lambda_i(1-\beta_i)\int_t^{\infty} e^{-(\lambda_i+\rho)(s-t)}\Big(\gamma_i \ln(\alpha_i X_i^*) - \varphi_i P\Big)\, ds, \quad i = 1, 2. \tag{21}$$

We obtain a system of equations, one for each country. In such system we need find two different but with the same form value functions V_i^{nc}, with its respective derivative $V_i^{nc'}$. We guess two identical functions, but for their economic parameters $A_i^{nc}P + B_i^{nc}$.

After the calculations enclosed in Appendix B, we find the non-cooperative X_i^{nc*}, A_i^{nc} and Ω_i. Note that we do not focus on solving B_i^{nc} in this paper, since it does not affect to the optimal emissions rule, nor the stock of pollution, in our linear-state differential game. Thus, the emissions rules

$$X_i^{nc^*} = -\frac{\gamma_i}{\omega_i V_i^{nc'}} = -\frac{\gamma_i}{\omega_i A_i^{nc}}, \quad i = 1, 2. \tag{22}$$

in which

$$A_i^{nc} = \frac{\varphi_i}{\rho + \delta} \times \Big[\Omega_i - 1\Big], \quad \Omega_i(\lambda_i, \beta_i),$$

with the following Ω_i terms, one for each country i

$$\Omega_i = \frac{\lambda_i(1 - \beta_i)}{\lambda_i + \rho + \delta}. \tag{23}$$

For the non-cooperative case, we find two different Ω_i, which depend not only on the common and fixed parameters ρ and δ, but also on the stochastic hyperbolic ones λ_i and β_i. Further, the two non-cooperative Ω_i affect in a very similar way as in the single-agent's Ω. Hence, for $\lambda_i = 0$ or $\beta_i = 1$, Ω_i cancels, and for $\lambda_i \to \infty$, the whole term becomes $1 - \beta_i$.

Proposition 2. Using the sophisticated agent procedure we find the following Nash-equilibria constant emissions rule and pollution stock trajectory.

The optimal emissions rules, for $i = 1,\ 2$, are given by

$$X_i^{nc} = -\frac{\gamma_i(\rho + \delta)}{\omega_i \varphi_i[\Omega_i - 1]}, \tag{24}$$

which are always positive due to the form of equation (23). No matter the values calibrated for parameters λ_i, β_i, ρ and δ, the value of Ω_i is always between [0, 1), which cancels the minus multiplier of equation (24). Note that $\lambda_i \in [0, \infty)$, $\beta_i \in [0, 1]$ and $\lambda_i + \rho + \delta \geq 0$. So that, $\lambda_i(1 - \beta_i) < \lambda_i + \rho + \delta$.

Furthermore, the time-consistent stock of pollution trajectory for non-cooperative sophisticated countries is

$$P^{nc}(s) = \left[P_t - \frac{\rho+\delta}{\delta} \cdot \sum_{i=1}^{2} \frac{-\gamma_i}{\varphi_i(\Omega_i - 1)}\right] e^{-\delta(s-t)} + \frac{\rho+\delta}{\delta} \cdot \sum_{i=1}^{2} \frac{-\gamma_i}{\varphi_i(\Omega_i - 1)}\,. \tag{25}$$

In a similar way to the single-agent case, if $\lambda_i = 0$, the term Ω_i will become 0, and we get the standard case stock

$$P^{nc}(s) = \left[P_t - \frac{\rho+\delta}{\delta} \cdot \sum_{i=1}^{2} \frac{\gamma_i}{\varphi_i}\right] e^{-\delta(s-t)} + \frac{\rho+\delta}{\delta} \cdot \sum_{i=1}^{2} \frac{\gamma_i}{\varphi_i}.$$

For the instantaneous gratification discounting, a similar result to the above equation is found. With the addition of β_i to the denominator, the multiplier becomes: $\frac{\gamma_i}{\varphi_i \beta_i}$, which increases the stock of pollution, $P(s)$, a lot when $s \to \infty$.

4.3 Cooperative case

In the cooperative game, both countries joint their individual functions J_i to maximize the sum of them J^c as one. It is reasonable to assume that they agree on their stochastic hyperbolic parameters, such the duration of the present interval τ exponentially distributed by a common λ. Even though they might face a different bias to the present β_i, we also assume they agree on a single β to avoid future extra deviations. In addition to the previous presented models, for the cooperative case each country has to be weighted by μ_i, where $0 \leq \mu_i \leq 1$, so that we can consider differences between countries. It is noteworthy to mention that countries in this setting are willing to cooperate due to the higher total output of the cooperative maximization compared to the sum of the individual's non-cooperative maximization results, $\sum_{i=1}^{n}(J_i^{nc}) \leq J^c$ for many cases of μ_i.

Therefore, the joint dynamic net payoff function to be maximized for cooperative agents with functional forms of the kind of equation (2) is characterized by

$$
\begin{aligned}
J^c &= E\bigg[\int_t^{t+\tau} e^{-\rho(s-t)} \sum_{i=1}^{2} \Big(\mu_i\gamma_i \ln(\alpha_i X_i) - \mu_i\varphi_i P\Big)\, ds \\
&+ \beta \int_{t+\tau}^{\infty} e^{-\rho(s-t)} \sum_{i=1}^{2} \Big(\mu_i\gamma_i \ln(\alpha_i X_i) - \mu_i\varphi_i P\Big)\, ds\bigg], \qquad (26)
\end{aligned}
$$

$$
s.t. \quad \dot{P} = \omega_1 X_1 + \omega_2 X_2 - \delta P, \quad P(t) = P_t. \qquad (27)
$$

Note that the dynamics of the stock of pollution are the same for non-cooperative and cooperative cases.

Following the same procedure as in the previous two cases, we describe the dynamic programming equation for cooperative agents with stochastic hyperbolic discounting. In this case, we have a single proposed cooperative value function V^c, which is the sum of two smaller V_1^c and V_2^c. In which case the dynamic programming equation is

$$
\rho V^c(P) + K^c(P) = \max_{X_i \geq 0} \left[\sum_{i=1}^{2} \Big(\mu_i\gamma_i \ln(\alpha_i X_i) - \mu_i\varphi_i P\Big) + V^{c'}(\omega_1 X_1 + \omega_2 X_2 - \delta P)\right], \qquad (28)
$$

where

$$
K^c(P) = \lambda(1-\beta) \int_t^{\infty} e^{-(\lambda+\rho)(s-t)} \sum_{i=1}^{2} \Big(\mu_i\gamma_i \ln(\alpha_i X_i^*) - \mu_i\varphi_i P\Big)\, ds. \qquad (29)
$$

The single-agent's case and the cooperative one are very similar. Thus, the joint net payoff function of the cooperative case $\sum_{i=1}^{n} \Big(\mu_i W_i(X_i, P)\Big)$ fills the same room as the net payoff function of the single-agent's case $W(X,P)$ in the dynamic programming equation. Nonetheless, the equation of motion, function $g(X_1, X_2, P)$, is equal to the non-cooperative case one, since countries accumulate emissions in the stock in the same way.

We also enclose all the calculations for the cooperative case in the Appendix C. We get two different emissions rules, one for each country, that depend to the common A^c, which is subject to each country's economic parameters and the selected weights. The emissions rules

$$
X_i^{c^*} = -\frac{\mu_i\gamma_i}{\omega_i V^{c'}} = -\frac{\mu_i\gamma_i}{\omega_i A^c}, \quad i = 1, 2; \qquad (30)
$$

in which

$$A^c = \sum_{i=1}^{2} \frac{\mu_i \varphi_i}{\rho+\delta} \times \Big[\Omega - 1\Big], \quad \Omega(\lambda, \beta),$$

with a single, similar to the one-agent case, term

$$\Omega = \frac{\lambda(1-\beta)}{\lambda+\rho+\delta}. \tag{31}$$

Again, the Ω term is always postive, no matter the values of its parameters, so as in previous cases.

Proposition 3. Using the sophisticated agent procedure we find the following time-consistent cooperative constant emissions rates, which only depends on each agent parameters, and time-consistent pollution stock trajectory.

The cooperative optimal emissions rules, for $i = 1, 2$, are determined by

$$X_i^c = -\frac{\rho+\delta}{\Omega-1} \times \frac{\mu_i \gamma_i}{\omega_i \sum_{i=1}^{2}(\mu_i \varphi_i)}\,. \tag{32}$$

The form of the emissions rules for the cooperative case are almost even to the non-cooperative one form, but for the sum of the parameters φ_i and the selected weights. In fact, in case those weights are equal to 1, $\mu_1 = \mu_2 = 1$, and also the parameters φ_i are equal, $\varphi_1 = \varphi_2$, the optimal emissions X_i^c are the half of X_i^{nc}.

Furthermore, the stock of pollution trajectory for cooperative countries is

$$P^c(s) = \left[P_t - \frac{\rho+\delta}{\delta(\Omega-1)} \cdot \frac{-\sum_{i=1}^{2}(\mu_i \gamma_i)}{\sum_{i=1}^{2}(\mu_i \varphi_i)}\right] e^{-\delta(s-t)} + \frac{\rho+\delta}{\delta(\Omega-1)} \cdot \frac{-\sum_{i=1}^{2}(\mu_i \gamma_i)}{\sum_{i=1}^{2}(\mu_i \varphi_i)}\,, \tag{33}$$

which, in the same way as in the two previous cases, if $\lambda = 0$ or $\beta = 1$, we will recover the standard case

$$P^c(s) = \left[P_t - \frac{\rho+\delta}{\delta} \cdot \frac{\sum_{i=1}^{2}(\mu_i \gamma_i)}{\sum_{i=1}^{2}(\mu_i \varphi_i)}\right] e^{-\delta(s-t)} + \frac{\rho+\delta}{\delta} \cdot \frac{\sum_{i=1}^{2}(\mu_i \gamma_i)}{\sum_{i=1}^{2}(\mu_i \varphi_i)}\,.$$

5 Sensitivity Analysis

In this section, we complement our previous results with a sensitivity analysis of the stochastic hyperbolic discounting parameters ($\beta - \lambda$) and discuss other parameters such as ($\gamma - \omega$). We first analyze the impact of such parameters to the optimum emissions rules for our three particular cases: *single-agent, non-cooperative* and *cooperative*. After doing so, we also analyze the effect of those parameters on the time-consistent stock of pollution trajectory. Moreover, we discuss the limits of our analysis and proposals for future research in the field.

We are interested in comparing the impact of λ for given values of β. To compute so, we follow the calibration for the stochastic hyperbolic parameters proposed by Zou et al. (2014), plus the calibration for scale economy parameters proposed by Jørgensen et al. (2003). Hence, the chosen values for the parameters are the following:

$$\rho = 0.046,\ \beta \in [0,1],\ \ \lambda \in [0,\infty),\ \ \gamma = (0.3, 0.4, 0.5,),\ \ \alpha = 1,\ \ \varphi = 0.2,$$
$$\omega = (0.2, 0.35),\ \ \delta = 0.3,\ \ \mu_1 = \mu_2 = 1;$$

where, we present four different cases of values of λ; we first discuss β and then we focus of its extreme cases; we compare whether the effects of having higher values of γ are similar to the experienced by having lower values of ω; and we set the weights of the two countries equal to 1, which allows us to compare the non-cooperative and cooperative frameworks.

5.1 Time-consistent emissions

Firstly, we examine how emissions evolve when the single-agent faces four different values of λ. Those emissions are constant with respect to time s, thus, we display their behavior when the parameter β changes. Doing so, we are able to distinguish and focus on the most important issues, regarding to the impact of λ, when discussing the emissions for non-cooperative and cooperative agents. The four λ different are displayed by (red, black, blue, green) normal lines.

Figure 1 shows that, when the single-agent faces a standard problem, $\lambda = 0$, the value of β does not affect emissions at all. It also shows that the case in which $\beta \to 1$, the impact of the chosen λ is negligible. In the case in which β is close to 0, which means that the agent is highly impatient, special attention must be focused on the instantaneous gratification case, in which time-consistent emissions are much higher than in any other case.

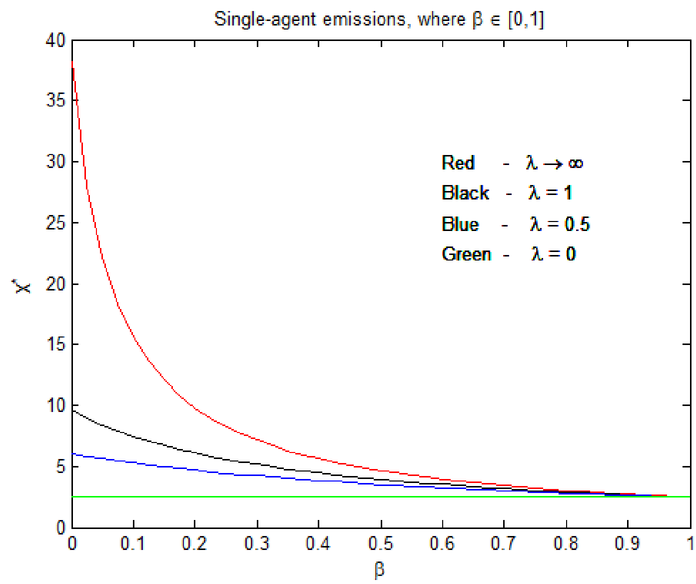

Figure 1: Impact of β and λ on X^*

We analyze now how, given the same values for λ and β, an increase in γ, which represents the single country's overall technology, has a similar impact on emissions as a decrease in ω, i.e., the single country's environmental-oriented technology. This is highly important for our non-cooperative and cooperative analysis, since ω does not affect the stock of pollution, which we study in the following subsection, due to the form of our studied linear-state differential game. Interestingly, for future research, a linear-quadratic differential game could solve this feature, since in Jørgensen et al. (2005) the authors show how the stock of pollution is affected by ω in a problem with standard discounting.

Figure 2 clearly shows that an increase in γ has a similar impact to a decrease in ω. Hence, we use different values of γ for countries 1 and 2 when analyzing the problem for non-cooperative and cooperative, as a proxy of the impact of different values of ω, the environmental technology.

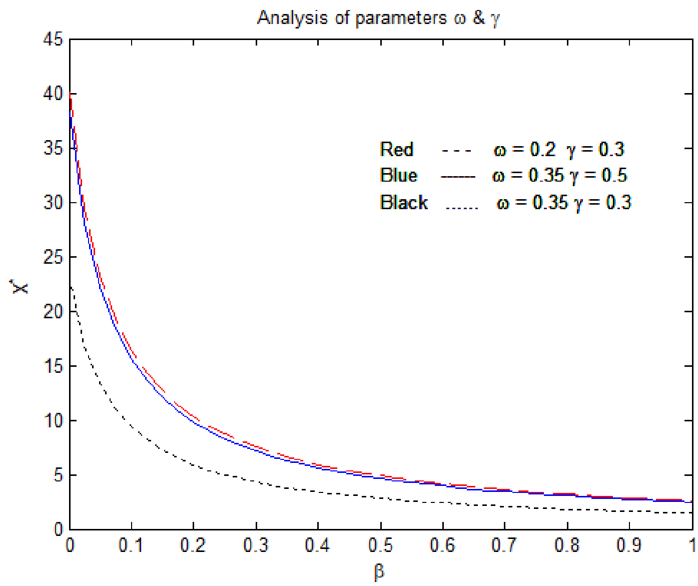

Figure 2: Sensitivity of ω and γ

We next show how emissions evolve for two different countries $\gamma = (0.5, 0.4)$ when facing the problem in the non-cooperative and cooperative way. We observe that for those countries with a higher value of γ, which can be interpreted as having a lower value of ω, i.e., less environmental-oriented technology, it is more worthy to have high emissions rates.

Figure 3 shows that, given both countries' weights are equal to one, the non-cooperative agents (red) are polluting twice as much as their respective cooperative agents (blue). This result is relatively insignificant when $\beta \to 1$. However, in the case in which $\beta \to 0$, doubling the emissions might be a serious issue. Moreover, there are also differences between those with higher and lower environmental-oriented technology as shown by dashed and normal lines respectively.

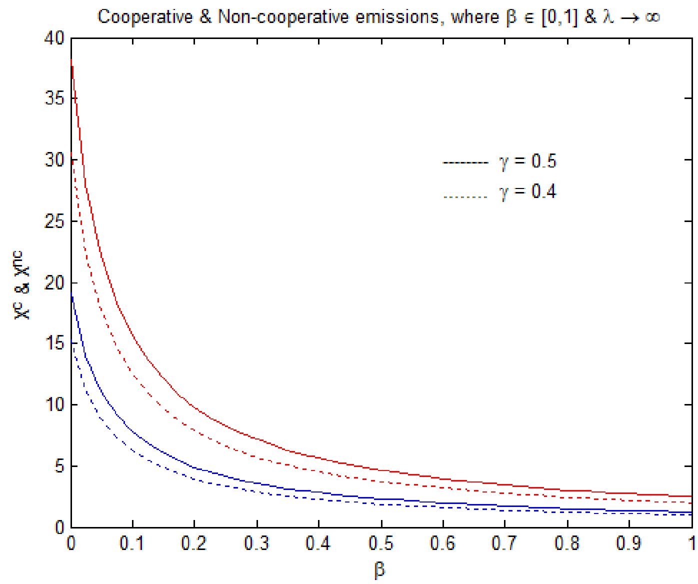

Figure 3: Cooperative vs Non-cooperative emissions

5.2 Time-consistent stock of pollution

We now analyze whether different values of λ have an impact on the stock of pollution when the parameter β is fixed. As we have observed in the previous subsection, the two most notable cases are when β is equal to 0 and 1. Thus, we build a two-graph figure which illustrates how time and λ affect the stock in these two particular cases for the single-agent economy.

Figure 4 shows the cases for a patient (blue) and an impatient (red) economy. It is clear that, the value of λ is not important if the economy is highly patient, i.e., $\beta = 0.9$. In fact, if we set $\beta = 1$, all four curves will be exactly the same. The case in which the economy is highly impatient $\beta = 0.2$, however, is more interesting. We observe big differences in the stock of pollution curves for our four different cases, in particular between the two extreme cases, but also on the intermediate ones.

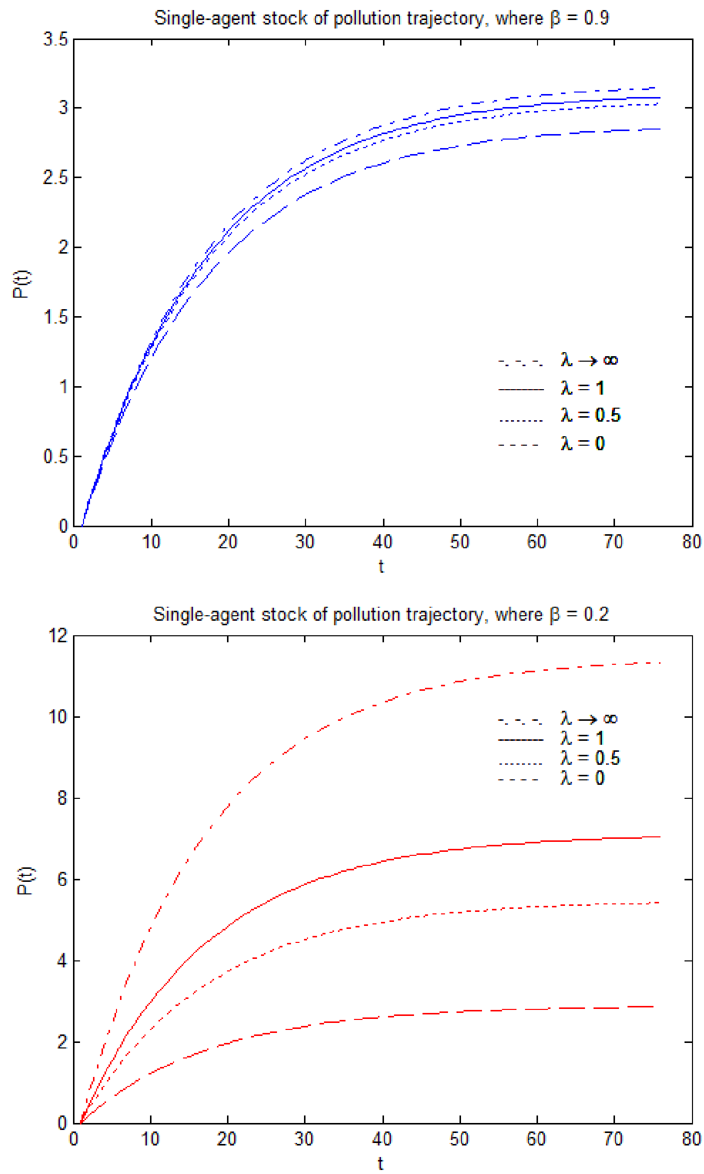

Figure 4: Impact of β and λ on the stock of pollution

Under these results, our last graph aims to study the dynamics of the cooperative and non-cooperative economies with a fixed high value for impatience, and the two most extreme cases for λ, standard and instantaneous gratification cases. We set the non-cooperative and cooperative economies cases in red and blue respectively, and separate them according to standard and instantaneous gratification as shown by the dashed and normal lines respectively. We also set $\gamma_1 = 0.5$ and $\gamma_2 = 0.4$ to illustrate the effect when those two countries have different technologies.

Figure 5 shows that, given a low value of β, the cases in which $\lambda = 0$ or $\lambda \to \infty$ have much higher effect, than the impact generated by cooperation or non-cooperation. Nonetheless, in a the same way as on Figure 3, the non-cooperative stock of pollution is at all times double that of the cooperative stock, which could imply severe consequences for those economies which are not willing to cooperate.

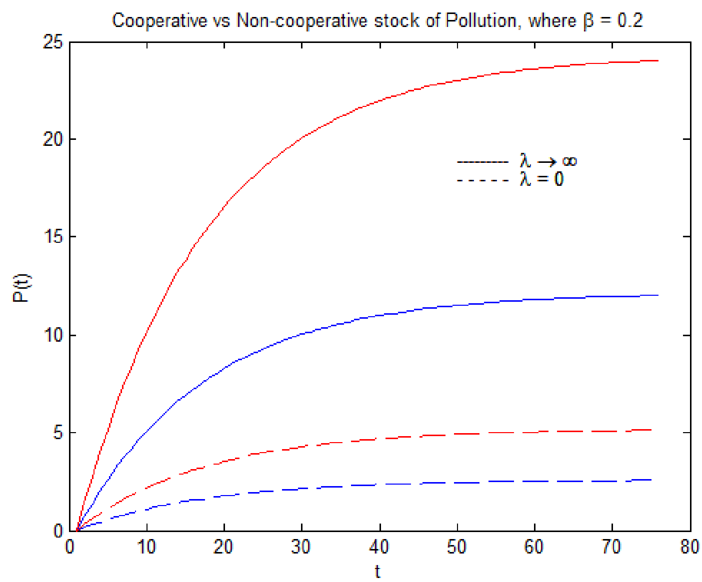

Figure 5: Cooperative vs Non-cooperative stock of pollution

Our results suggest that different values of λ imply very different outputs, and have a very important effect on the economy, mainly in the cases where β is low. Thus, economies should make the effort to track and calibrate their respective value of β_1 as well as the values of other regions β_i, $i = 2, ..., n$, which also have an indirect effect on them. In addition, a properly weighting μ_i is needed in order to study the real difference between cooperation and non-cooperation, which in our case is always double the amount of emissions and stock.

6 Concluding Remarks

We have computed the sophisticated-agent solution with stochastic hyperbolic discounting in infinite time horizon for a single-agent case, two non-cooperative agents case and two cooperative agents case, as an extension of Harris and Laibson (2013) and Zou et al. (2014). We have contributed to such existing literature by introducing this kind of discounting into the framework of dynamic games of the environment, such as Jørgensen et al. (2003). In addition, an exhaustive sensitivity analysis of the stochastic hyperbolic parameters has been conducted.

The discussion of this work might be useful for policy implementation towards transboundary pollution agreements between countries. Notwithstanding, future research should consider more representative payoff functions, such as the linear-quadratic function, which possibly, would have allowed emissions to be dependant with respect to time.

In all three studied cases, the parameter representing impatient behavior was trascendental, so that tracking and, perhaps, imposing restrictions to such parameter in real economies could be important to assure environmental quality in the long run. Furthermore, tracking the hazard rate by which economies' present interval is distributed, would be of a high interest.

On the one hand, we find that the time-consistent emission policies are constant and independent of time, only depending on the chosen scale parameters, although this result may change for other payoff functions. On the other hand, the time-consistent stock of pollution for all cases is an exponential function that can be highly impacted in some situations, such as the instantaneous gratification case. In addition, we find that, at all cases, non-cooperative agents pollute double that of equally to one weighted cooperative agents.

From our study, there is plenty room for extensions and improvements in new research. For instance, considering more realistic payoff functions, country agreements by incentive strategies or computing the precommitted-agent and naive-agent solutions.

Appendix A

Since for the single-agent case, the proposed value function is $V = A\,P + B$, and its derivative $V' = A$, by maximizing the right hand side of the dynamic programming equation (12) with respect to X, we find

$$\begin{aligned}
&0 = \frac{\gamma}{X} + \omega\,V'(P), \\
&\frac{X}{\gamma} = -\frac{1}{\omega\,V'(P)}, \\
&X^* = -\frac{\gamma}{\omega\,A}\,.
\end{aligned} \tag{A.1}$$

To find the time-consistent state of the stock of pollution we integrate the law of motion of pollution

$$\begin{aligned}
\dot{P} &= \omega X^* - \delta P\,; \\
\dot{P} + \delta P &= \omega X^* = \omega(-\frac{\gamma}{\omega\,A})\,.
\end{aligned}$$

The above first-order differential equation can be solved by the integrating factor method. Thus, we multiply the whole thing by $e^{\delta\tau}$ and integrate from t to s

$$\begin{aligned}
&\int_t^s e^{\delta\tau}(\dot{P} + \delta P)d\tau = \int_t^s e^{\delta\tau}\omega(-\frac{\gamma}{\omega\,A})d\tau\,, \\
&P(s)e^{\delta s} - P(t)e^{\delta t} = \omega(-\frac{\gamma}{\omega\,A})\int_t^s e^{\delta\tau}d\tau\,, \\
&P(s)e^{\delta s} - P(t)e^{\delta t} = \omega(-\frac{\gamma}{\omega\,A})\,\frac{e^{\delta s} - e^{\delta t}}{\delta}\,, \\
&P(s)e^{\delta s} = P(t)e^{\delta t} - \frac{\gamma}{A}\cdot\frac{1 - e^{-\delta(s-t)}}{\delta}e^{\delta s}\,.
\end{aligned}$$

Let $\Upsilon(s,t)$ be:

$$\Upsilon(s,t) = -\frac{\gamma}{A} \cdot \frac{1-e^{-\delta(s-t)}}{\delta} e^{\delta s}; \tag{A.2}$$

thus,

$$\begin{aligned}
e^{\delta s} P(s) &= P(t)\, e^{\delta t} + \Upsilon(s,t)\,, \\
P(s) &= \Big[P(t)\, e^{\delta t} + \Upsilon(s,t)\Big]\, e^{-\delta s}, \\
&= \Big[P_t\, e^{\delta t} - \frac{\gamma}{A} \cdot \frac{1-e^{-\delta(s-t)}}{\delta} e^{\delta s}\Big]\, e^{-\delta s}, \\
&= P_t\, e^{-\delta(s-t)} - \frac{\gamma}{A} \cdot \frac{1-e^{-\delta(s-t)}}{\delta}\,, \\
&= \Big[P_t\, - \frac{-\gamma}{\delta A}\Big]\, e^{-\delta(s-t)} - \frac{\gamma}{\delta A}\,; \qquad \text{(A.3)}
\end{aligned}$$

furthermore, if we fix $t = 0$ and $s = t$ the 0-agent rule can be computed

$$P(t) = \Big[P_0\, - \frac{-\gamma}{\delta A}\Big]\, e^{-\delta t} - \frac{\gamma}{\delta A}\,.$$

Using equation $(A.3)$, and replacing it into $K(P)$ - equation (13) - we are able to write all P in function of actual time t instead of s, which also allows split equation (A.4). Hence,

$$\begin{aligned}
K(P) &= \lambda(1-\beta) \int_t^{\infty} e^{-(\lambda+\rho)(s-t)} \Big(\gamma \ln(\alpha X^*) - \varphi P\Big)\, ds, \\
&= \lambda(1-\beta) \int_t^{\infty} e^{-(\lambda+\rho)(s-t)} \Big(\gamma \ln(\alpha X^*) - \varphi \Big[P(t)\, e^{\delta t} + \Upsilon(s,t)\Big]\, e^{-\delta s}\Big)\, ds, \\
&= \lambda(1-\beta) \Big[\int_t^{\infty} e^{-(\lambda+\rho+\delta)(s-t)} \Big(-\varphi P(t)\Big)\, ds \\
&\qquad + \int_t^{\infty} e^{-(\lambda+\rho)(s-t)} \Big(\gamma \ln(\alpha X^*) - \varphi \Upsilon(s,t)\, e^{-\delta s}\Big)\, ds\Big].
\end{aligned}$$

We use the above expression and replace it into the Dynamic programming equation (DPE), $K(P)$ together with $V(P)$ and V'. Further, we do not have to replace X^* by $-\frac{\gamma}{\omega A}$ yet, since P does not appear explicitly inside X^*. Thus,

$$\begin{aligned}
& \rho(AP_t) + \lambda(1-\beta)\int_t^\infty e^{-(\lambda+\rho+\delta)(s-t)}\Big(-\varphi P_t\Big)\,ds \\
+ \quad & \rho B + \lambda(1-\beta)\int_t^\infty e^{-(\lambda+\rho)(s-t)}\Big(\gamma\ln(\alpha X^*) - \varphi\Upsilon(s,t)\,e^{-\delta s}\Big)\,ds \\
= \quad & \gamma\ln(\alpha X^*) - \varphi P_t + \omega A X^* - \delta A P_t\,; \qquad \text{(A.4)}
\end{aligned}$$

where we suppose A constant and with the same value for all P_t. Under this assumption, we focus now in all terms from equation (A.4) multiplied by P_t to find A:

$$\rho(AP_t) + \lambda(1-\beta)\int_t^\infty e^{-(\lambda+\rho+\delta)(s-t)}\Big(-\varphi P_t\Big)\,ds = -\,\varphi P_t - \delta A P_t\,.$$

Then, we collect all P_t terms and isolate A

$$\begin{aligned}
&\rho A - \lambda(1-\beta)\varphi\int_t^\infty e^{-(\lambda+\rho+\delta)(s-t)}ds = -\,\varphi - \delta A\,, \\
&\rho A + \delta A = \lambda(1-\beta)\varphi\int_t^\infty e^{-(\lambda+\rho+\delta)(s-t)}ds - \varphi\,, \\
&A = \lambda(1-\beta)\frac{\varphi}{\rho+\delta}\int_t^\infty e^{-(\lambda+\rho+\delta)(s-t)}ds - \frac{\varphi}{\rho+\delta}\,, \\
&\quad = \lambda(1-\beta)\frac{\varphi}{\rho+\delta}\cdot\frac{1}{\lambda+\rho+\delta} - \frac{\varphi}{\rho+\delta}\,, \\
&\quad = \frac{\varphi}{\rho+\delta}\left[\frac{\lambda(1-\beta)}{\lambda+\rho+\delta} - 1\right];
\end{aligned}$$

let the term $\Omega = \frac{\lambda(1-\beta)}{\lambda+\rho+\delta}$. So that,

$$A = \frac{\varphi}{\rho+\delta}\times\Big[\Omega - 1\Big]\,. \qquad \text{(A.5)}$$

We find in equation $(A.5)$ a constant solution for A, which will depend only on the scale parameters and on selected (λ, β).

Now, we replace A from equation $(A.5)$ into equation $(A.1)$, allowing us to find the efficient emissions rate X^*. Which is

$$\begin{aligned} X^* &= -\frac{\gamma}{\omega \, \frac{\varphi}{\rho+\delta} \cdot [\Omega - 1]}\,, \\ &= -\frac{\gamma(\rho+\delta)}{\omega\, \varphi\, [\Omega - 1]}\,. \end{aligned} \tag{A.6}$$

Using the expressions $(A.3)$ and $(A.6)$ we find the equilibrium stock of pollution trajectory at time t

$$\begin{aligned} P(s) &= \left[P_t - \frac{-\gamma}{\delta \, \frac{\varphi}{\rho+\delta} \times [\Omega - 1]}\right] e^{-\delta(s-t)} - \frac{\gamma}{\delta \, \frac{\varphi}{\rho+\delta} \times [\Omega - 1]}\,, \\ &= \left[P_t - \frac{-\gamma(\rho+\delta)}{\delta\, \varphi[\Omega - 1]}\right] e^{-\delta(s-t)} - \frac{\gamma(\rho+\delta)}{\delta\, \varphi[\Omega - 1]}\,, \\ &= \left[P_t - \frac{(\rho+\delta)\gamma}{\delta\varphi} \times \frac{-1}{\Omega - 1}\right] e^{-\delta(s-t)} + \frac{(\rho+\delta)\gamma}{\delta\varphi} \times \frac{-1}{\Omega - 1}\,; \end{aligned} \tag{A.7}$$

we leave the term $\frac{-1}{\Omega-1}$ separated because it is equal to 1 for the general case when $\lambda = 0 \to \Omega = 0$. Thus, it is easier to compare between expressions $(A.8)$ and $(A.9)$. Given this,

$$P(s) = \left[P_t - \frac{(\rho+\delta)\gamma}{\delta\varphi}\right] e^{-\delta(s-t)} + \frac{(\rho+\delta)\gamma}{\delta\varphi}. \tag{A.8}$$

Appendix B

Our proposal for the value functions for non-cooperative agents are: $A_i^{nc}P + B_i^{nc}$ for $i = 1, 2$. According to that, by maximizing the right hand side of dynamic programming equation (20) with respect to X_i, we find

$$\begin{aligned}
0 &= \frac{\gamma_i}{X_i} + \omega_i\, V_i^{'}(P), \\
\frac{X_i}{\gamma_i} &= -\frac{1}{\omega_i\, V_i'(P)}, \\
X_i^* &= -\frac{\gamma_i}{\omega_i\, A_i^{nc}}\,, \quad i = 1, 2. \qquad \text{(B.1)}
\end{aligned}$$

To find the time-consistent state of the stock of pollution we compute the problem for the particular law of motion of pollution

$$\dot{P} = \omega_1 X_1^* + \omega_2 X_2^* - \delta P\,;$$

$$\dot{P} + \delta P = \sum_{i=1}^{2} \omega_i X_i^* = \sum_{i=1}^{2} \frac{-\gamma_i \omega_i}{\omega_i\, A_i^{nc}} = -\sum_{i=1}^{2} \frac{\gamma_i}{A_i^{nc}}\,;$$

The above first-order differential equation can be solved by the integrating factor method. Thus, we multiply the whole thing by $e^{\delta\tau}$ and integrate from t to s

$$\begin{aligned}
\int_t^s e^{\delta\tau}(\dot{P} + \delta P)d\tau &= -\int_t^s e^{\delta\tau} \sum_{i=1}^{2} \frac{\gamma_i}{A_i^{nc}} d\tau\,, \\
P(s)e^{\delta s} - P(t)e^{\delta t} &= -\sum_{i=1}^{2} \frac{\gamma_i}{A_i^{nc}} \int_t^s e^{\delta\tau} d\tau\,, \\
&= -\sum_{i=1}^{2} \frac{\gamma_i}{A_i^{nc}} \frac{e^{\delta s} - e^{\delta t}}{\delta}\,, \\
P(s)e^{\delta s} &= P(t)e^{\delta t} - \sum_{i=1}^{2} \frac{\gamma_i}{A_i^{nc}} \cdot \frac{1 - e^{-\delta(s-t)}}{\delta} e^{\delta s}\,.
\end{aligned}$$

Let $\Upsilon^{nc}(s,t)$ be:

$$\Upsilon^{nc}(s,t) = -\sum_{i=1}^{2} \frac{\gamma_i}{A_i^{nc}} \cdot \frac{1-e^{-\delta(s-t)}}{\delta} e^{\delta s}. \tag{B.2}$$

Thus, we compute the time s stock of pollution using the equation (B.2),

$$\begin{aligned}
P(s)\, e^{\delta s} &= P(t)\, e^{\delta t} + \Upsilon^{nc}(s,t)\,, \\
P(s) &= \Big[P(t)\, e^{\delta t} + \Upsilon^{nc}(s,t)\Big]\, e^{-\delta s} \\
&= \Big[P_t\, e^{\delta t} - \sum_{i=1}^{2} \frac{\gamma_i}{A_i^{nc}} \cdot \frac{1-e^{-\delta(s-t)}}{\delta} e^{\delta s}\Big]\, e^{-\delta s} \\
&= \Big[P_t - \sum_{i=1}^{2} \frac{-\gamma_i}{\delta A_i^{nc}}\Big]\, e^{-\delta(s-t)} - \sum_{i=1}^{2} \frac{\gamma_i}{\delta A_i^{nc}}\,. \qquad \text{(B.3)}
\end{aligned}$$

Notably, we can easily compute the 0-agent rule by fixing $t = 0$ and $s = t$,

$$P(t) = \Big[P_0 - \sum_{i=1}^{2} \frac{-\gamma_i}{\delta A_i^{nc}}\Big]\, e^{-\delta t} - \sum_{i=1}^{2} \frac{\gamma_i}{\delta A_i^{nc}}\,.$$

Using equation $(B.3)$, and replacing it into $K_i(P)$, equation (21), we are able to write all P in function of actual time t instead of s, which also allows split equation (B.4) and find A_i^{nc}, equation (B.5)

$$\begin{aligned}
K_i(P) &= \lambda_i(1-\beta_i) \int_t^{\infty} e^{-(\lambda_i+\rho)(s-t)} \Big(\gamma_i \ln(\alpha_i X_i^*) - \varphi_i P\Big)\, ds, \\
&= \lambda_i(1-\beta_i) \int_t^{\infty} e^{-(\lambda_i+\rho)(s-t)} \Big(\gamma_i \ln(\alpha_i X_i^*) - \varphi_i \Big[P(t)\, e^{\delta t} + \Upsilon^{nc}(s,t)\Big]\, e^{-\delta s}\Big)\, ds, \\
&= \lambda_i(1-\beta_i) \Big[\int_t^{\infty} e^{-(\lambda_i+\rho+\delta)(s-t)} \Big(-\varphi_i P(t)\Big)\, ds \\
&\qquad + \int_t^{\infty} e^{-(\lambda_i+\rho)(s-t)} \Big(\gamma_i \ln(\alpha_i X_i^*) - \varphi_i \Upsilon^{nc}(s,t)\, e^{-\delta s}\Big)\, ds\Big].
\end{aligned}$$

We will use the above expression and replace it into the DPE, $K_i(P)$ together with $V_i(P)$ and V_i'. Further, we do not have to replace X_i^* by $-\frac{\gamma_i}{\omega_i A_i^{nc}}$ yet, since it does not depend on P.

$$\begin{aligned}
& \rho(A_i^{nc}P_t) + \lambda_i(1-\beta_i)\int_t^{\infty} e^{-(\lambda_i+\rho+\delta)(s-t)}\Big(-\varphi_i P_t\Big)\,ds \\
+ \quad & \rho B_i^{nc} + \lambda_i(1-\beta_i)\int_t^{\infty} e^{-(\lambda_i+\rho)(s-t)}\Big(\gamma_i \ln(\alpha_i X_i^*) - \varphi_i \Upsilon^{nc}(s,t)\, e^{-\delta s}\Big)\,ds \\
= \quad & \gamma_i \ln(\alpha_i X_i^*) - \varphi_i P_t + A_i^{nc}\sum_{i=1}^{2}(\omega_i X_i^*) - \delta A_i^{nc} P_t \,.Thus,
\end{aligned} \tag{B.4}$$

We suppose A_i^{nc} constant and with the same value for all P_t. We focus now in all the terms multiplied by P_t to find A_i^{nc},

$$\rho(A_i^{nc}P_t) + \lambda_i(1-\beta_i)\int_t^{\infty} e^{-(\lambda_i+\rho+\delta)(s-t)}\Big(-\varphi_i P_t\Big)\,ds = -\,\varphi_i P_t - \delta A_i^{nc} P_t\,,$$

we collect all P_t terms,

$$\begin{aligned}
& \rho A_i^{nc} - \lambda_i(1-\beta_i)\varphi_i \int_t^{\infty} e^{-(\lambda_i+\rho+\delta)(s-t)} ds = -\,\varphi_i - \delta A_i^{nc}\,, \\
& \rho A_i^{nc} + \delta A_i^{nc} = \lambda_i(1-\beta_i)\varphi_i \int_t^{\infty} e^{-(\lambda_i+\rho+\delta)(s-t)} ds - \varphi_i\,, \\
& A_i^{nc} = \lambda_i(1-\beta_i)\frac{\varphi_i}{\rho+\delta}\int_t^{\infty} e^{-(\lambda_i+\rho+\delta)(s-t)} ds - \frac{\varphi_i}{\rho+\delta}\,, \\
& \quad\; = \lambda_i(1-\beta_i)\frac{\varphi_i}{\rho+\delta}\cdot\frac{1}{\lambda_i+\rho+\delta} - \frac{\varphi_i}{\rho+\delta}\,, \\
& \quad\; = \frac{\varphi_i}{\rho+\delta}\left[\frac{\lambda_i(1-\beta_i)}{\lambda_i+\rho+\delta} - 1\right],
\end{aligned}$$

where, let the terms $\Omega_i = \frac{\lambda_i(1-\beta_i)}{\lambda_i+\rho+\delta}$, such

$$A_i^{nc} = \frac{\varphi_i}{\rho+\delta} \times \Big[\Omega_i - 1\Big]\,. \tag{B.5}$$

We find in equation $(B.5)$ a constant solution for A_i^{nc}, which will depend only on the scale parameters and on the selected λ_i.

Now, we replace A_i^{nc} from equation $(B.6)$ into equation $(B.1)$, allowing us to find the efficient emissions rate

$$
\begin{aligned}
X_i^{nc} &= -\frac{\gamma_i}{\omega_i \frac{\varphi_i}{\rho+\delta} \cdot [\Omega_i - 1]} \,, \\
&= -\frac{\gamma_i(\rho+\delta)}{\omega_i \, \varphi_i \, [\Omega_i - 1]} \,.
\end{aligned}
\tag{B.6}
$$

Using the expressions $(B.3)$ and $(B.6)$ we find the equilibrium stock of pollution trajectory at time t

$$
\begin{aligned}
P^{nc}(s) &= \left[P_t - \sum_{i=1}^{2} \frac{-\gamma_i}{\delta \frac{\varphi_i}{\rho+\delta} \cdot [\Omega_i - 1]}\right] e^{-\delta(s-t)} - \sum_{i=1}^{2} \frac{\gamma_i}{\delta \frac{\varphi_i}{\rho+\delta} \cdot [\Omega_i - 1]} \,, \\
&= \left[P_t - \frac{\rho+\delta}{\delta} \sum_{i=1}^{2} \frac{-\gamma_i}{\varphi_i[\Omega_i - 1]}\right] e^{-\delta(s-t)} - \frac{\rho+\delta}{\delta} \sum_{i=1}^{2} \frac{\gamma_i}{\varphi_i[\Omega_i - 1]} \,, \\
&= \left[P_t - \frac{\rho+\delta}{\delta} \sum_{i=1}^{2} \frac{-\gamma_i}{\varphi_i(\Omega_i - 1)}\right] e^{-\delta(s-t)} + \frac{\rho+\delta}{\delta} \sum_{i=1}^{2} \frac{-\gamma_i}{\varphi_i(\Omega_i - 1)} \,.
\end{aligned}
\tag{B.7}
$$

We leave the term $\frac{-1}{\Omega_i - 1}$ separated because it is equal to 1 for the standard case when $\lambda = 0 \rightarrow \Omega_i = 0$. Thus, it is easier to compare between expressions $(B.7)$ and $(B.9)$

$$
P^{nc}(s) = \left[P_t - \frac{\rho+\delta}{\delta} \sum_{i=1}^{2} \frac{\gamma_i}{\varphi_i}\right] e^{-\delta(s-t)} + \frac{\rho+\delta}{\delta} \sum_{i=1}^{2} \frac{\gamma_i}{\varphi_i}. \tag{B.8}
$$

Appendix C

By maximizing the right hand side of dynamic programming equation (29) with respect to X_i and using our proposed value function for the cooperative case $V^c(P)$, we find

$$
\begin{aligned}
0 &= \frac{\mu_i\gamma_i}{X_i^*} + \omega_i\, V'(P),\\
\frac{X_i^*}{\mu_i\gamma_i} &= -\frac{1}{\omega_i\, V'(P)},\\
X_i^c &= -\frac{\mu_i\gamma_i}{\omega_i\, A^c}\,. \qquad \text{(C.1)}
\end{aligned}
$$

To find the time-consistent state of the stock of pollution we consider the law of motion of pollution

$$\dot{P} = \omega_1 X_1^c + \omega_2 X_2^c - \delta P\,;$$

$$\dot{P} + \delta P = \sum_{i=1}^{2} \omega_i X_i^c = \sum_{i=1}^{2} \frac{-\mu_i\gamma_i\omega_i}{\omega_i\, A^c} = -\sum_{i=1}^{2} \frac{\mu_i\gamma_i}{A^c}\,.$$

The above first-order differential equation can be solved by the integrating factor method. Thus, we multiply the whole thing by $e^{\delta\tau}$ and integrate from t to s

$$
\begin{aligned}
\int_t^s e^{\delta\tau}(\dot{P} + \delta P)d\tau &= -\int_t^s e^{\delta\tau} \sum_{i=1}^{2} \frac{\mu_i\gamma_i}{A^c} d\tau\,,\\
P(s)e^{\delta s} - P(t)e^{\delta t} &= -\sum_{i=1}^{2} \frac{\mu_i\gamma_i}{A^c} \int_t^s e^{\delta\tau} d\tau\,,\\
&= -\sum_{i=1}^{2} \frac{\mu_i\gamma_i}{A^c} \cdot \frac{e^{\delta s} - e^{\delta t}}{\delta}\,,\\
P(s)e^{\delta s} &= P(t)e^{\delta t} - \sum_{i=1}^{2} \frac{\mu_i\gamma_i}{A^c} \cdot \frac{1 - e^{-\delta(s-t)}}{\delta} e^{\delta s}\,.
\end{aligned}
$$

Let $\Upsilon^c(s,t)$ be:

$$\Upsilon^c(s,t) = -\sum_{i=1}^{2} \frac{\mu_i \gamma_i}{A^c} \cdot \frac{1 - e^{-\delta(s-t)}}{\delta} e^{\delta s}. \tag{C.2}$$

Hence, using equation (C.2) inside $P(s)$ we build the stock of pollution trajectory

$$\begin{aligned}
P(s)\, e^{\delta s} &= P(t)\, e^{\delta t} + \Upsilon^c(s,t)\,, \\
P(s) &= \Big[P(t)\, e^{\delta t} + \Upsilon^c(s,t)\Big]\, e^{-\delta s}, \\
&= \Big[P_t\, e^{\delta t} - \sum_{i=1}^{2} \frac{\mu_i \gamma_i}{A^c} \cdot \frac{1 - e^{-\delta(s-t)}}{\delta} e^{\delta s}\Big]\, e^{-\delta s}, \\
&= P_t\, e^{\delta(s-t)} - \sum_{i=1}^{2} \frac{\mu_i \gamma_i}{A^c} \cdot \frac{1 - e^{-\delta(s-t)}}{\delta}, \\
&= \Big[P_t\, - \frac{-1}{\delta A^c} \sum_{i=1}^{2} (\mu_i \gamma_i)\Big]\, e^{-\delta(s-t)} + \frac{-1}{\delta A^c} \sum_{i=1}^{2} (\mu_i \gamma_i)\,. \qquad (C.3)
\end{aligned}$$

In which, we obtain the 0-agent stock of pollution trajectory by setting: $t = 0$ and $s = t$,

$$P(t) = \Big[P_0\, - \frac{-1}{\delta A^c} \sum_{i=1}^{2} (\mu_i \gamma_i)\Big]\, e^{-\delta t} + \frac{-1}{\delta A^c} \sum_{i=1}^{2} (\mu_i \gamma_i)\,.$$

Using equation $(C.3)$ and replacing it into $K^c(P)$, equation (29), we are able to write all P in function of actual time t instead of s, which also allows us to split equation (C.4) and, so that, obtain equation (C.5). Thus,

$$\begin{aligned}
K_i(P) &= \lambda(1-\beta) \int_t^{\infty} e^{-(\lambda+\rho)(s-t)} \sum_{i=1}^{2} \Big(\mu_i \gamma_i \ln(\alpha_i X_i^c) - \mu_i \varphi_i P\Big)\, ds, \\
&= \lambda(1-\beta) \int_t^{\infty} e^{-(\lambda+\rho)(s-t)} \sum_{i=1}^{2} \Big(\mu_i \gamma_i \ln(\alpha_i X_i^c) - \mu_i \varphi_i \Big[P(t)\, e^{\delta t} + \Upsilon^c(s,t)\Big]\, e^{-\delta s}\Big)\, ds, \\
&= \lambda(1-\beta) \Big[\int_t^{\infty} e^{-(\lambda+\rho+\delta)(s-t)} \sum_{i=1}^{2} \Big(-\mu_i \varphi_i P(t)\Big)\, ds \\
&\qquad + \int_t^{\infty} e^{-(\lambda+\rho)(s-t)} \sum_{i=1}^{2} \Big(\mu_i \gamma_i \ln(\alpha_i X_i^c) - \mu_i \varphi_i \Upsilon^c(s,t)\, e^{-\delta s}\Big)\, ds\Big].
\end{aligned}$$

We use the above expression and replace it into the dynamic programming equation $K^c(P)$ together with $V^c(P)$ and $V^{c'}$. Further, we do not have to replace X_i^c by $-\frac{\mu_i\gamma_i}{\omega_i A^c}$ yet, since it does not depend on P. Given this,

$$\begin{aligned} & \rho(A^c P_t) + \lambda(1-\beta)\int_t^\infty e^{-(\lambda+\rho+\delta)(s-t)}\Big(-\mu_i\varphi_i P_t\Big)\,ds \\ + \quad & \rho B^c + \lambda(1-\beta)\int_t^\infty e^{-(\lambda+\rho)(s-t)}\Big(\mu_i\gamma_i\ln(\alpha_i X_i^c) - \mu_i\varphi_i\Upsilon^c(s,t)\,e^{-\delta s}\Big)\,ds \\ = \quad & \sum_{i=1}^{2}\Big(\mu_i\gamma_i\ln(\alpha_i X_i^c) -\ \mu_i\varphi_i P_t\Big) + A^c\sum_{i=1}^{2}(\omega_i X_i^c) - \delta A^c P_t\,. \end{aligned} \tag{C.4}$$

We suppose A^c constant and with the same value for all P_t. Under this assumption, we focus now in all the terms multiplied by P_t to find A^c

$$\rho(A^c P_t) + \lambda(1-\beta)\int_t^\infty e^{-(\lambda+\rho+\delta)(s-t)}\sum_{i=1}^{2}\Big(-\mu_i\varphi_i P_t\Big)\,ds = -\sum_{i=1}^{2}\mu_i\varphi_i P_t - \delta A^c P_t\,,$$

collect all P_t terms and solve for A^c,

$$\begin{aligned} & \rho A^c - \lambda(1-\beta)\sum_{i=1}^{2}(\mu_i\varphi_i)\int_t^\infty e^{-(\lambda+\rho+\delta)(s-t)}ds = -\sum_{i=1}^{2}(\mu_i\varphi_i) - \delta A^c\,, \\ & \rho A^c + \delta A^c = \lambda(1-\beta)\sum_{i=1}^{2}(\mu_i\varphi_i)\int_t^\infty e^{-(\lambda+\rho+\delta)(s-t)}ds - \sum_{i=1}^{2}(\mu_i\varphi_i)\,, \\ & A^c = \lambda(1-\beta)\sum_{i=1}^{2}\frac{\mu_i\varphi_i}{\rho+\delta}\int_t^\infty e^{-(\lambda+\rho+\delta)(s-t)}ds - \sum_{i=1}^{2}\frac{\mu_i\varphi_i}{\rho+\delta}\,, \\ & \quad = \lambda(1-\beta)\sum_{i=1}^{2}\frac{\mu_i\varphi_i}{\rho+\delta}\cdot\frac{1}{\lambda+\rho+\delta} - \sum_{i=1}^{2}\frac{\mu_i\varphi_i}{\rho+\delta}\,, \\ & \quad = \sum_{i=1}^{2}\frac{\mu_i\varphi_i}{\rho+\delta}\cdot\left[\frac{\lambda(1-\beta)}{\lambda+\rho+\delta} - 1\right]. \end{aligned}$$

Let $\Omega = \frac{\lambda(1-\beta)}{\lambda+\rho+\delta}$, then

$$A^c = \sum_{i=1}^{2}\frac{\mu_i\varphi_i}{\rho+\delta}\times\Big[\Omega - 1\Big]\,, \tag{C.5}$$

where we find a constant solution for A^c, which will depend only on the scale parameters and on the selected λ.

Now, we replace A^c from equation $(C.5)$ into equation $(C.1)$, allowing us to find the efficient emissions rates

$$
\begin{aligned}
X_i^c &= -\frac{\mu_i \gamma_i}{\omega_i \sum_{i=1}^2 \frac{\mu_i \varphi_i}{\rho+\delta} \cdot [\Omega - 1]}\,, \\
&= -\frac{\mu_i \gamma_i (\rho + \delta)}{\omega_i \, [\Omega - 1] \sum_{i=1}^2 (\mu_i \varphi_i)}\,, \\
&= -\frac{\rho + \delta}{\Omega - 1} \cdot \frac{\mu_i \gamma_i}{\omega_i \sum_{i=1}^2 (\mu_i \varphi_i)}, \quad i = 1, 2. \qquad \text{(C.6)}
\end{aligned}
$$

Using the expressions $(C.3)$ and $(C.6)$ we compute the time-consistent stock of pollution trajectory

$$
\begin{aligned}
P^c(s) &= \left[P_t - \frac{-1}{\delta \sum_{i=1}^2 \frac{\mu_i \varphi_i}{\rho+\delta} \cdot [\Omega - 1]} \sum_{i=1}^2 (\mu_i \gamma_i) \right] e^{-\delta(s-t)} + \frac{-1}{\delta \sum_{i=1}^2 \frac{\mu_i \varphi_i}{\rho+\delta} \cdot [\Omega - 1]} \sum_{i=1}^2 (\mu_i \gamma_i)\,, \\
&= \left[P_t - \frac{\rho + \delta}{\delta(\Omega - 1)} \cdot \frac{-\sum_{i=1}^2 (\mu_i \gamma_i)}{\sum_{i=1}^2 (\mu_i \varphi_i)} \right] e^{-\delta(s-t)} + \frac{\rho + \delta}{\delta(\Omega - 1)} \cdot \frac{-\sum_{i=1}^2 (\mu_i \gamma_i)}{\sum_{i=1}^2 (\mu_i \varphi_i)}\,, \qquad \text{(C.7)}
\end{aligned}
$$

where we leave the term $\frac{-1}{\Omega-1}$ separated because it is equal to 1 when, $\lambda = 0 \rightarrow \Omega = 0$. Thus, it is easier to compare between expressions $(C.7)$ and $(C.8)$, which illustrates the standard case trajectory

$$
P^c(s) = \left[P_t - \frac{\rho + \delta}{\delta} \cdot \frac{\sum_{i=1}^2 (\mu_i \gamma_i)}{\sum_{i=1}^2 (\mu_i \varphi_i)} \right] e^{-\delta(s-t)} + \frac{\rho + \delta}{\delta} \cdot \frac{\sum_{i=1}^2 (\mu_i \gamma_i)}{\sum_{i=1}^2 (\mu_i \varphi_i)}\,. \qquad \text{(C.8)}
$$

References

Ainslie, G. and N. Haslam, 1992. Hyperbolic discounting. *Choice Over Time, Russell Sage, New York.*

Barro, R. J., 1999. Ramsey meets Laibson in the neoclassical growth model. *The Quarterly Journal of Economics* 114 (4): 1125 – 1152.

Dockner, E. and N. Van Long, 1993. International pollution control: Cooperative versus non-cooperative strategies. *Journal of Environmental and Management* 24: 13 – 29.

Harris, C. and D. Laibson, 2013. Instantaneous Gratification. *Quarterly Journal of Economics* 1999: 205 – 248.

Havely, Y., 2015. Time Consistency: Stationarity and Time Invariance. *Econometrica* 83 (1): 335 – 352.

Jørgensen, S., Martín-Herrán, G. and G. Zaccour, 2003. Agreeability and time consistency in linear-state differential games. *Journal of Optimization Theory and Applications* 119 (1): 31 – 39.

Jørgensen, S., Martín-Herrán, G. and G. Zaccour, 2005. Sustainability of cooperation over time in linear-quadratic differential games. *International Game Theory Review* 7 (4): 395 - 406.

Karlan, D., McConnell, M., Mullainathan, S. and J. Zinnan, 2016. Getting to the top of Mind: How Reminders Increase Saving. *Management Science* 2016: 3393 – 3411.

Karp, L., 2005. Global warming and hyperbolic discounting. *Journal of Public Economics* 89: 261 - 282.

Laibson, D., 1997. Golden Eggs and Hyperbolic Discounting. *The Quarterly Journal of Economics* 112 (2): 443 – 477.

Loewenstein, G. F. and D. Prelec, 1992. Anomalies in Intertemporal choice: Evidence and an interpretation. *The Quarterly Journal of Economics* 107 (2): 573 – 597.

Marín-Solano, J. and J. Navas, 2009. Non-constant discounting in finite horizon: the terminal time case. *Journal of Economic Dynamics and Control* 33 (3): 666 – 675.

Marín-Solano, J. and J. Navas, 2010. Consumption and portfolio rules for time-inconsistent investors. *European Journal of Operational Research* 201 (3): 860 – 872.

McClure, S. M., Laibson, D., Lowenstein, G. and J.D. Cohen, 2004. Separate neutral systems value immediate and delayed monetary rewards. *Science* 306 (5695): 503 – 507.

Merton, R. C., 1969. Lifetime portfolio selection under uncertainty: the continuous-time case. *The review of Economics and Statistics* 51 (3): 247 – 257.

Merton, R. C., 1971. Optimum consumption and portfolio rules in a continuous-time model. *Journal of Economic Theory* 3 (4): 373 – 413.

O'Donoghue, T and M. Rabin, 1999. Doing It Now or Later. *The American Economic Review* 103 – 124.

Petrosjan, L. A. and G. Zaccour, 2003. Time-consistent Shapley value of pollution cost reduction. *Journal of Economics Dynamics and Control* 27 (3): 381 – 398.

Samuelson, P.A., 1937. A Note on Measurement of Utility. *The Review of Economic Studies* 4 (2): 155 – 161.

Strotz, R. H., 1956. Myopia and Inconsistency in Dynamic Utility Maximization. *The Review of Economic Studies* 23 (3): 165 – 180.

Thaler, R. H., 1981. Some empirical evidence on dynamic inconsistency. *Economic Letters* 8: 201 – 207.

Thaler, R. H. and H. M. Shefrin, 1981. An Economic Theory of Self-Control. *Journal of Political Economy* 89(2): 392 – 406.

van der Ploeg, F. and A. de Zeeuw, 1992. International aspects of pollution control. *Environmental and Resource Economics* 2 (2): 117 – 139.

Zou, Z., Chen, S. and L. Wedge, 2014. Finite horizon consumption and portfolio decisions with stochastic hyperbolic discounting. *Journal of Mathematical Economics* 52: 70 – 80.